5-2017

A Methodology for Identifying Sources of Disparities in the Socio-Economic Impacts of ICT Capabilities in Sub-Saharan Economies

Sergey Samoilenko
Averett University, ssamoilenko@averett.edu

Kweku-Muata Osei-Bryson
Virginia Commonwealth University, KMOsei@vcu.edu

A METHODOLOGY FOR IDENTIFYING SOURCES OF DISPARITIES IN THE SOCIO-ECONOMIC IMPACTS OF ICT CAPABILITIES IN SUB-SAHARAN ECONOMIES

Sergey Samoilenko
Department of CS/CIS
Averett University, USA
ssamoilenko@averett.edu

Kweku-Muata Osei-Bryson
Department of Information Systems
Virginia Commonwealth University, USA
KMOsei@VCU.Edu

Abstract

There is general agreement that Information and Communication Technologies (ICT) may deliver transformative socio-economic impacts. However, there is no general agreement on the mechanisms by which the impacts are delivered, for purely efficiency-driven assessment of the impacts leaves many important factors out. To further inquire into the issue of the impact of ICT capabilities, as well as into the factors that may play a role in delivering the impact of such capabilities, we develop and test a methodology for obtaining insights into the context-specific mechanisms by which ICT capabilities are translated into socio-economic outcomes. The methodology is tested in the context of 24 economies of Sub-Saharan Africa.

Keywords

ICT Capabilities, Sub-Saharan Economies, Data Envelopment Analysis, Market Basket Analysis, Decision Trees Induction

1. Introduction

At this point it is expected that the impact of Information and Communication Technologies (ICT) at the country level extends beyond purely economic gains (e.g., via growth in productivity) and into the sphere of social development (Eide, 2015). While wealthier economies of the world may look towards optimization of the economic impact of ICT, poorer countries of Sub-Saharan Africa (SSA) should be in a position of reaping a transformational-level of socio-economic benefits of ICT. It is hard to determine whether the transformational impact within the context of SSA is indeed taking place, but we could start the assessment by investigating the link between the state of ICT and its socio-economic impacts. The premise is that for a sustained transformational impact to take place an economy needs to obtain and maintain an efficient path of transforming ICT capabilities into socio-economic outcomes.

Benchmarking is one of the tools by which improvements in efficiency could be obtained, and we suggest that this tool could be utilized by SSA economies to improve the performance of their ICT capabilities. However, it is a bridge too far for SSA economies to benchmark developed countries outright, for the disparity in the levels of accumulated and developed ICT infrastructure and annual investments is too great to disregard. Thus, we suggest that as a first step we investigate the efficiency of socio-economic impact of ICT capabilities within a group of SSA. Such investigation would entail identification of the economies that are more efficient in obtaining socio-economic benefits of ICT, and then proceeding with identifying the characteristics of such economies vis-à-vis characteristics of the less efficient economies, all within the context of SSA.

We feel that such inquiry is well-justified, because "…the complex relationships between ICTs and socioeconomic performance are not fully understood and their causality not fully established" (Di Battista, Dutta, Geiger & Lanvin, 2015, p.4). Consequently, this problem presents an important research opportunity to investigate and we formulate the mail goal of this study as follows:

The purpose of this investigation is to identify some of the factors that differentiate groups of Sub-Saharan economies in regard to their levels of wealth and efficiency of socio-economic impact of ICT.

This would require accomplishing the following two research objectives:

To develop a methodology for identifying a set of group-specific characteristics of economies reflecting their state of economic development and efficiency of socio-economic impacts of ICT.
To apply the developed methodology within the context of Sub-Saharan economies to identify factors associated with group-specific disparities in economic development and socio-economic impact of ICT.

Resultantly, by conducting this investigation we aim to contribute to the existing body of knowledge in the area of ICT4D in more than one way. First, we develop a methodology allowing for identifying a combination of characteristics describing various groups of SSA. While Decision Trees analysis could be performed to identify the factors specific to various groups of economies (Samoilenko and Osei-Bryson, 2014), this technique is not well suited for identifying *combinations* of factors. In this study we demonstrate how Association Rule Mining (ARM) and Decision Trees Induction (DTI) could be used *in synergy* to identify a set of attributes differentiating various groups of SSA economies. The true novelty of this study is that our approach allows for *identifying sets of attributes based on differentiating factors.* Second, while the previous inquiries concentrated either on economic or on social impacts of ICT capabilities, our study aims to be more comprehensive in this regard- we investigate both social and economic impact of ICT capabilities within the same sample of SSA economies. Finally, the results of empirical analysis should provide valuable information to policy and decision makers working in the area of ICT4D within the context of SSA.

To summarize, accomplishing our research objectives would contribute to the current state of knowledge in the area of ICT for development in several ways including:

A development of a novel three-phase methodology for identifying a set of rules and differentiating factors that, taken together, allow for gaining deeper insights in disparities between the groups of economies.
An identification of the factors and complex associations impacting the disparity of the economic development, socio-economic impact of ICT, as well as of the efficiency of the impact in the context of SSA

We conduct our investigation within the context of 27 economies of SSA, using the data set for the period of 2011-2014. The analysis of the data is supported by a three-phase methodology utilizing Data Envelopment Analysis (DEA), Decision Trees Induction (DTI), and Association Rule Mining (ARM).

2. Research Framework

In our investigation we rely on the framework of Networked Readiness (Dutta, Geiger & Lanvin, 2015), the adapted version of which is depicted below in Figure 1. The framework relies on four subindexes and their ten sub-categories (or *pillars*) to obtain the value of the Networked Readiness Index (NRI), which reflects the capacity of economies to benefit from ICT. An increase in the value of NRI for a given economy is indicative of the increase of the impact of ICT on innovation and productivity (Dutta & Jain, 2003). Interestingly, the original framework does not explicitly connect *Environment*, *Readiness*, and *Usage* subindexes (referred to as *Drivers* within the framework) with *Impact* subindex (referred to as *Impact*), despite relying on a principle that "…the environment, readiness, and use—interact, co-evolve, and reinforce each other to create greater impact" (Di Battista, Dutta, Geiger & Lanvin, 2015, p.4).

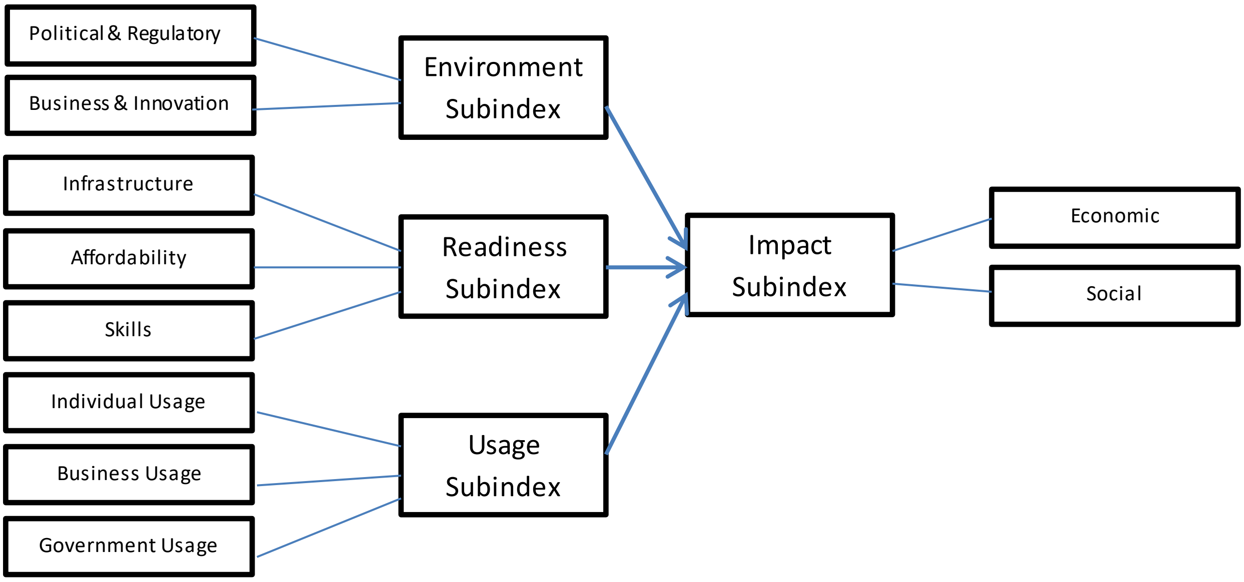

Figure 1 The Framework of Networked Readiness, Adapted

We scope our inquiry by only considering relationships between *Drivers* (environment, readiness, and use) and *Impact* (socio-economic impact of *Drivers*) - as indicated by arrows in Figure 1.All possible interactions within *Drivers* (e.g., between environment, readiness, and use) we consider to be beyond the scope of our investigation. In our inquiry we use the framework depicted in Figure 1 to investigate, via DEA, the efficiency of the process by which *Environment*, *Readiness*, and *Usage* subindexes impact two subcategories of *Impact* subindex-*Economic* and *Social* impacts. Once we identify the better and worse performers of the sample, we utilize DTI to identify some of the factors that differentiate groups of economies in regard to their level of economic development and efficiency of socioeconomic impact of ICT. Then we use ARM, via *Market Basket Analysis*, to attempt to identify a set of rules describing the groups.

3. Proposed Methodology

Our proposed methodology consists of three phases where each phase involves the application of well established data analysis methods (i.e. Data Envelopment Analysis (DEA), Decision Tree Induction (DTI), and Association Rule Mining (ARM). Prior offering a description of the phases of our methodology we would like to explain to our reader what makes our approach truly novel.

3.1 A New Methodology: Benefits and Justifications

DEA is a method that is widely used for the purposes of calculating scores of the relative efficiency of entities that receive inputs and produce outputs. For example, we could compare three groups of basketball players of different levels (e.g., high school, college, and professional) in terms of their efficiency of conversion of minutes played and attempts taken into assists and points. Results of DEA would yield the most relatively efficient group, but because DEA model is a "black box" model we would not know what differentiates the groups, or why one group is more efficient than the other two.

The insights could be provided by DTI , which would yield an attribute, or a few attributes, that differentiate the groups. Thus, given a set of attributes describing the three groups of basketball players we may find out that the main difference between the groups is in years of experience and hours of weekly practice. However, there are plenty of players with many years of experience who train long hours every week, but, nevertheless, don't play so well. We would like to know what set of attributes is actually associated with the outputs- assists or points scored. Here is where ARM may help.

ARM allows for generating sets of association rules of the type "if (a,b,c) then (d)". This is very valuable, for we can see the patterns of associations specific to each group. However, ARM tends to generate many rules, some trivial, some meaningless/non-actionable, and some useful. The problem with selecting the rules describing the different group is that the rules may contain completely different attributes- this would result in comparing apples and oranges. For example, in the case of basketball players we may get "if (height > x) then (minutes_played > y" for high school players, "if (experience > n) then (assists > m)" and so on. So, the trick is to identify a set of rules that is based on a set of common criteria that differentiates the groups- this insight is provided by DTI.

Consequently, the novelty of our approach is associated with its capability to identify the main differentiating factors responsible for heterogeneity of the context, and then to base the selection of the rules on those factors.

Previous investigations used a hybrid DEA/DTI methodology (Samoilenko and Osei-Bryson, 2007), and the use of ARM with DEA was recently reported by Samoilenko (2016); however, this investigation represents the first case of using the three methods (i.e. DEA, DTI, ARM) *in synergy*. Simply put, if DEA allows us to identify the efficient performers, and DTI helps us to discover the relevant dimensions that differentiate efficient and inefficient performers, then ARM allows us to benchmark efficient performers via a set of "*IF THEN*" rules that rely on the discovered by DTI dimensions. To our knowledge, no other combination of data analytic and data mining methods could offer so much in so few steps.

3.2 Phase 1: Data Envelopment Analysis (DEA)

During the first phase we rely on DEA to evaluate relative efficiency of three "*Drivers* $\rightarrow$ *Impact*" paths. We will use *variable return to scale* (VRS) DEA model to conduct the analysis, for it is reasonable to argue that SSA economies have not yet reached the point of developing a level of ICT infrastructure allowing accruing the benefits yielded by capitalizing on economies of scale.

Given a four-year time period we will run DEA 12 times. Consequently, for each economy in the sample we are going to have four scores of relative efficiency for each of the three models. At this point we need to provide a justification for the inputs and outputs included in our models. In regard to outputs the reasoning is intuitive- first, we would like to assess the efficiency of the overall impact, and then, each type of the impact separately. This is because an economy could be efficient in obtaining one type of an impact (e.g., economic) and not efficient in regard to another impact (e.g., social).

DEA Model	Inputs of DEA Model	Outputs of DEA Model
$Drivers \rightarrow$ Overall_*Impact* (DOI)	Environment Subindex Readiness Subindex Usage Subindex	Impact Subindex
$Drivers \rightarrow$ Economic_*Impact* (DEI)	Environment Subindex Readiness Subindex Usage Subindex	Economic Impact Sub-category
$Drivers \rightarrow$ Social_*Impact* (DSI)	Environment Subindex Readiness Subindex Usage Subindex	Social Impact Sub-category

Table 2 DEA models of the study

With regard to the choice of the inputs of DEA model, our approach is methodological. While we are free to use eight sub-categories of Drivers as inputs of a DEA model, the general rule of thumb is that for a reasonable level of discrimination number of economies (or *Decision Making Units* in DEA terms) must be at least twice the product of inputs and outputs (Dyson, Allen, Camanho, Podinovski, Sarrico, & Shale, 2001). In our case we have a sufficient number of economies in our set, but if we use a DEA model with eight inputs and two outputs then we would need to have at least 2*8*2 = 32 economies in the sample.

Furthermore, and more importantly, the greater the number of factors included in the DEA model, the lower the level of discrimination of the model (Dyson et al., 2001). However, we would like to use all the data available to us so we could inquire, for example, whether a set of specific factors- pillars- differentiate relatively efficient economies from relatively inefficient once. We would use DTI to do so.

Additionally, we use DEA to calculate the values of the Malmquist index (MI) - this allows us to assess the changes in relative efficiency of SSA economies that took place over time. Such results would not only identify the economies that exhibited growth in productivity (under assumption of constant return to scale), but to also identify the sources of growth (EC-change in efficiency vs. TC- change in technology). By applying DTI we can identify factors that differentiate the *Growth vs. No Growth* economies.

3.3 Phase 2: Decision Tree Induction (DTI)

To proceed with Phase 2 we need to create a new variable "*Target*" to differentiate various groups of economies. We are interested in three types of groupings: first, we would like to differentiate the groups of SSA by their level of economic development, and then we would like to differentiate the groups in terms of their efficiency of the socioeconomic impact of ICT. The last analysis would involve differentiating SSA economies by growth in productivity – *Growth vs. No Growth*. Thus, we would conduct DTI three times, which would require Target to have three domains of values.

In the first case, grouping by income, the domain of values of *Target* would be {1, 2, 3}, for, respectively, Low Income (LI), Low Middle Income (LM), and Upper Middle (UM) groups of economies. In the second case, grouping by efficiency, Target would assume the values of {0, 1}, for, respectively, relatively inefficient, and relatively efficient SSA. The same domain of values, namely, {0, 1}, could be applied to the grouping by growth in productivity, where "*0*" would indicate "*No Growth*" and "*1*" would indicate "*Growth*."

3.4 Phase 3: Association Rule Mining (ARM)

The purpose of Phase 3 is to find possible patterns, associations, or causal structures that may exist in our data. One of the main advantages of ARM is that it is suitable for undirected data mining; thus, we'll aim to discover naturally occurring associations between the factors (sub-indexes of *Drivers* and *Impact*) - components of NRI. ARM could be classified as either being explanatory or exploratory in nature. In the case of our investigation we employ exploratory ARM, for we do not have any theoretical support for why certain relationships between the sub-indexes of NRI should exist. A very common approach to generating associations between the variables, or *itemsets*, via ARM is by using the *apriori* algorithm (Agrawal & Ramakrishnan, 1994) - we will rely on this approach in the current investigation. Transformation of the data is required for this step- we follow the method of Samoilenko (2016) to do so.

4. Research Questions and Null Hypotheses of the Study

At this point we can operationalize the two objectives of this investigation in the form of the specific research questions and corresponding null hypotheses.

The first research question operationalizes the first objective as follows:

Is the developed methodology capable of generating sets of differentiating factors and association rules for a given set of criteria?

One of these criteria is associated with the level of income of economies (e.g., Low Income vs. Low-Middle vs. Upper-Middle), while another criterion is a relative level of efficiency (e.g., relatively efficient vs. relatively inefficient) of *Drivers* $\rightarrow$ *Impact* path, and the third one is a growth in productivity that took place over period of time.

We can answer this research question by testing the corresponding null hypotheses:

H01a: The DTI part of the methodology will fail to generate a set of differentiating factors characterized by high-level splits.
We will test H01a under the conditions of: *high-level splits that differentiate at least 60% of at least one of the groups of SSA economies.*
H01b: The ARM part of the methodology will fail to generate sets of association rules for a given set of criteria.

We will test H01b under the minimal conditions of *Support* > 20%, *Confidence* > 1.0, and *Lift* > 1.0.

While the results of DTI and ARM may offer useful insights by themselves, we would like to use the two methods in a complementary fashion; thus, we state another hypothesis as follows:

H01c: The results of DTI and ARM are not complementary.

We will test H01c under the condition that the differentiating factors identified by DTI would be included in sets of rules identified by ARM.

The second research question operationalizes the second objective as follows:

Does the choice of criteria such as level of economic development, relative efficiency, and growth in productivity impact the combination of factors describing various groups of economies and relationships between Drivers and Impacts of ICT?

Basically, we would like to find out if the different criteria could be characterized via different set of factors- this allows us to inquire into the specificity of a setting expressed as a combination of sub-categories of NRI. We will answer the second research question after testing our second null hypothesis:

H02: No combination of factors contained in the generated association rules would be unique to a given context.

The simple side-by-side comparison of the generated association rules and split variables will serve as a sufficient criterion for testing H02.

5. The Data

We obtained the data from a reputable source- the World Economic Forum's Global Information Technology Report 2015 (GITR, 2015). In 2012 the representation of NRI was partially changed in terms of the number and representation of the pillars of three sub-indexes of NRI; it was also the year when the *Impact* subindex was introduced. Given the changes that took place between 2011 and 2012, we decided to concentrate on the new version of NRI and collect the data provided in GITR 2012, 2013, 2014, and 2015. In some cases the representation of SSA economies was inconsistent- for example, we could not include Angola, Seychelles, Liberia, Gabon, Sierra Leone, and Guinea in our sample because the data for some of the years was missing.

While there is an advantage to increasing the sample size of a study, there is a price to pay via dealing with missing variables, imputation of values, and additional data preprocessing. After considering the pluses and minuses of "sample size vs. data *actually* available" we have assembled a smaller data set that contained no missing data and no outliers, but was as reliable as one could get from a given source.

Income Level	**Sub-Saharan Economies**
Low Income	Burkina Faso, Burundi, Chad, Ethiopia, Gambia, Kenya, Madagascar, Malawi, Mali, Mozambique, Rwanda, Tanzania, Uganda, Zimbabwe
Low- Middle Income	Cameroon, Cape Verde, Côte d'Ivoire, Ghana, Lesotho, Nigeria, Senegal, Swaziland, Zambia
Upper- Middle Income	Botswana, Mauritius, Namibia, South Africa

Table 1 Sample of Sub-Saharan Economies, by Income Level

Overall, we were able to compile the data set representing 27 economies of Sub-Saharan Africa (the classification of the International Monetary Fund as of October 2014). The sample consists of 14 low income economies, nine low-middle economies, and four upper-middle economies (the classification of the World Bank as of July 2014). Membership of the each group of the sample is provided in Table 1.

6. Results of the Data Analysis

6.1 Phase 1: Data Envelopment Analysis (DEA)

We offer a summary of the results of DEA below. If a given economy has been determined to be relatively efficient for at least three times over the period of four years, we have labeled such economy as "*efficient*" for the whole period of four years. Because our economies fall within three distinct groups- low income (LI), low-middle income (LM), and upper-middle income (UM), we also determined the relative efficiency of each economy over the four years within its group – we will use this information in Phase 3 when we perform ARM**.**

Our results demonstrated that seven economies out of the full sample are relatively efficient with regard to the impact of *Drivers* on social, economic, and overall *Impact* of ICT. Additionally, we identified relatively efficient economies per each of the income-level group; in some cases (e.g., Burundi, Chad, Kenya, Mali, Rwanda, and Senegal) the relatively efficient within its group' economies are also efficient overall. In other cases (e.g., Swaziland, Lesotho, Botswana, Mauritius, Namibia, and South African Republic) the relatively inefficient, overall, economies end up being efficient within their respective group.

Income Level	**Economy**	**Overall Efficiency of the impact of ICT**	**Overall Changes in Productivity,**	**Growth via EC?**	**Growth via TC?**
LI	BFA	Efficient	Growth	Yes	Yes
LI	BDI	Efficient	Growth	Yes	No
LI	TCD	Efficient	No growth	No	No
LI	ETH	Efficient	No growth	No	No
LI	GMB	Inefficient	Growth	No	No
LI	KEN	Efficient	No growth	No	No
LI	MDG	Inefficient	Growth	Yes	No
LI	MWI	Inefficient	No growth	No	No
LI	MLI	Efficient	Growth	No	Yes
LI	MOZ	Inefficient	No growth	Yes	No
LI	RWA	Efficient	No growth	Yes	No
LI	TZA	Inefficient	Growth	Yes	No
LI	UGA	Inefficient	No growth	Yes	No
LI	ZWE	Inefficient	Growth	Yes	No
LM	CMR	Inefficient	Growth	Yes	No
LM	CPV	Inefficient	No growth	No	No
LM	GHA	Inefficient	No growth	No	No
LM	NGA	Inefficient	No growth	No	No
LM	SEN	Efficient	No growth	No	Yes
LM	SWZ	Inefficient	Growth	Yes	No
LM	ZMB	Inefficient	Growth	Yes	No
LM	CIV	Inefficient	Growth	Yes	No
LM	LSO	Inefficient	Growth	Yes	No
UM	BWA	Inefficient	No growth	No	No
UM	MUS	Inefficient	Growth	Yes	No
UM	NAM	Inefficient	No growth	Yes	No

UM	ZAF	Inefficient	No growth	No	No

Table 4 Results of DEA

Additionally, we used DEA to calculate the values of Malmquist index (MI), which allows us to assess the changes in the scores of relative efficiency that took over period of time. Under the assumption of constant returns to scale the change indicates changes in productivity. Consequently, we are able to assess whether the economies become more productive or not. Because MI is comprised of two components- change in efficiency (EC) and change in technology (TC), we are also able to assess whether the changes in productivity are associated with a particular component. Overall, only 11 economies (40% of the sample) exhibited growth in productivity.

An analysis of the changes in EC and TC offers an interesting insight: 16 economies (60% of the sample) exhibited positive changes in efficiency, but only 3 economies (11% of the sample) demonstrated positive changes in technology. Finally, it is worth noting that only one economy, Burkina Faso, exhibited a balanced growth in productivity, when the growth was driven by both components of MI. Overall, the picture suggests that SSA, as a group, would benefit from a better technology- this suggests that investments in ITC infrastructure should be prioritized.

Finally, it is worth noting that 8 economies (30% of the sample) have not only exhibited a decline in productivity, but have exhibited a decline in terms of both components of MI.

Overall, we summarize the results of DEA part of our methodology as follows:

Seven economies out of the full sample are relatively efficient with regard to the impact of *Drivers* on social, economic, and overall *Impact* of ICT. There are relatively efficient economies per each of the income-level group. Only 11 economies (40% of the sample) exhibited growth in productivity, while 8 economies (30% of the sample) have exhibited a decline SSA, as a group, would benefit from a better technology.

6.2 Phase 2: Decision Tree Induction (DTI)

Grouping by	Group	Differentiating/Split Variable
Economic Development	Low Income vs. Low Middle Income vs. Upper Middle Income	Individual Usage Business Usage Skills Readiness
Relative Efficiency	Relatively Efficient vs. Relatively Inefficient	Individual Usage Economic Impact
Change in Productivity	Growth vs. No Growth	Infrastructure Readiness Affordability Readiness Individual Usage Social Impact

Table 5 Results of Decision Trees Analysis

The results of DTI allow us to test our first null hypothesis, H01a, for decision tree induction did generate high-level splits that differentiated groups of economies. Results summarized in the table above show that such pillars of NRI as *Individual Usage*, *Business Usage*, and *Skills Readiness* do play important role in differentiating three groups of economies. It is not surprising that there is appear to be a clear-cut difference between Low Income and Upper-

Middle Income economies, and much less of a difference between Low-Middle Income economies and the other two.

We could also identify *Individual Usage* and *Economic Impact* as pillars that play role in differentiating relatively efficient SSA economies from inefficient ones. It appears that Infrastructure Readiness, Affordability Readiness, and Individual Usage are factors playing role in differentiating those economies that became more productive from those that didn't.

At this point, we summarize the results of DTI part of our methodology as follows:

Pillars of NRI such as *Individual Usage*, *Business Usage*, and *Skills Readiness* do play important role in differentiating three groups of economies

Individual Usage and *Economic Impact* are pillars that play a role in differentiating relatively efficient SSA economies from inefficient ones

Infrastructure Readiness, Affordability Readiness, and Individual Usage differentiate those economies that became more productive from those that didn't.

6.3 Phase 3: Association Rule Mining (ARM)

The results summarized in Table 6 allow us to test our null hypotheses. First, the results allow us to reject *H01b*, for the application of ARM did result in the generation of multiple association rules under the criteria of *Support* > 20%, *Confidence* > 1.0, and *Lift* > 1.0. Second, the results also allow us to reject *H02*, for the generated by ARM rules contain context-specific combinations of factors.

Condition	Generated Rules			Sup.	Conf.	Lift
Low Income	low IND_USE, low BUS_USE	=>	low ECON_IMP	21%	0.7	1.9
	low SKILL_READ, low BUS_USE	=>	low ECON_IMP	21%	0.6	1.5
Low-Middle	midhigh BUS_USE	=>	midlow SOCIO_IMP	20%	0.5	2.5
	midhigh SKILL_READ	=>	midhigh SOCIO_IMP	25%	0.5	1.5
Upper-Middle	high SKIL_READ, high AFFORD_READ, high GOV_USE	=>	high SOCIO_IMP	32%	1.0	2.3
	high BUS&INNOV_ENV, high INFR_READ, high BUS_USE	=>	high ECON_IMP	31%	0.85	2.3
	high BUS&INNOV_ENV, high IND_USE, high BUS_USE	=>	high ECON_IMP	31%	0.85	2.2
Low Income, Inefficient	low INFR_READ, low BUS_USE	=>	low SKILL_READ	36%	1.0	1.9
	low IND_USE, low BUS_USE	=>	low SKILL_READ	32%	1.0	1.9

Low-Middle, Efficient	low BUS_USE, low GOV_USE	=>	low ECON_IMP	25%	1.0	4.0
	low BUS_USE, low ECON_IMP	=>	low SOCIO_IMP	25%	1.0	3.0

Table 6 Impact-Specific Rules for Low Income and Low-Middle Income economies

Finally, the results summarized in Table 6 allow us to reject *H01c* that *the results of DTI and ARM are not complementary. DTI identified Individual Usage, Business Usage*, and *Skills Readiness as the variables differentiating the groups of economies in our sample.*

7. Discussion of the Results

The results of the data analysis presented in the previous sections offer evidence that we were successful in addressing the research questions of this study. We developed and tested a novel methodology allowing for investigating complex context-specific relationships between the factors reflecting the state and the impact of ICT capabilities. The discussion of the results is presented along the points that we considered noteworthy.

First,
Despite the presence of complex relationships between the Drivers and Impacts of ICT there are common themes associated with the levels of the scores of factors comprising NRI-Business Usage, Individual Usage, and Skills Readiness appear to have a direct relationship with the levels of the scores of socio-economic Impact of ICT.
We point out that while the variety of association rules has been generated for a different set of criteria, a common line could also be glanced- some subcategories of NRI' subindexes (e.g., related to Skills, Business, Individual usage)appear more frequently than other subcategories.

Second,
Results of our investigation suggest that Business Usage and Individual Usage are among the factors that appear to differentiate economies in terms of their level of economic development, as well as in terms of their relative efficiency of the impact of ICT on the socioeconomic bottom line.

These results suggest that wealthier and more efficient economies tend to have higher scores of *Business Usage and Individual Usage*. The presence of a simple association between the level of income of an economy, its efficiency, and ICT usage seems to be apparent.

Third,
Infrastructure and Affordability of ICT seem to have an impact on growth in productivity of SSA economies.

This finding is important because it was provided by two different methods of analysis- DEA and DTI. According to the results of DEA only 3 economies exhibited growth in technology over the period 2012- 2015, and DTI independently confirmed it by identifying *Infrastructure* and *Affordability* as factors differentiating growth vs. no growth economies of SSA.

8. **Conclusion**

In this investigation we developed and applied a methodology allowing for generating sets of association rules from the combination of factors describing relationships between *Drivers* and *Impact* of ICT. The results of the data analysis do confirm the notion that the relationships between the factors representing *Drivers* and *Impact* are indeed complex. However, the underlying complexity of the relationship could be made more transparent to researchers and practitioners by the developed in this study methodology.

References

Agrawal, R. and Ramakrishnan S. (1994). Fast Algorithms for Mining Association Rules in Large Databases. In Proceedings of the 20th International Conference on Very Large Data Bases (VLDB '94), Jorge B. Bocca, Matthias Jarke, and Carlo Zaniolo (Eds.). Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 487-499.

Eide, E.B. (2015). Preface, in Global Information Technology Report 2015, p. v.

Di Battista, A., Dutta, S., Geiger, T., & Lanvin, B. (2015). The Networked Readiness Index 2015: Taking the Pulse of the ICT Revolution, Global Information Technology Report 2015, 3-28.

Dutta, S. & Jain, A. (2003). The Networked Readiness of Nations, in eds. Dutta, S., A. Lanvin, B., & Paua, F. The Global Information Technology Report 2002–2003, New York, Oxford, Oxford University Press.

Dutta, S., Geiger, T. & Lanvin, B. (Eds.) (2015). Global Information Technology Report 2015. ICTs for Inclusive Growth. Geneva: World Economic Forum and INSEAD. Retrieved April 20, 2015 from http://www3.weforum.org/docs/WEF_Global_IT_Report_2015.pdf

Dyson, R.G., Allen R., Camanho A.S., Podinovski, V.V., Sarrico C.S., & Shale, E.A. (2001). Pitfalls and Protocols in DEA, European Journal of Operational Research, Vol. 132, pp. 245-259.

GITR. (2015). World Economic Forum' Global Information Technology Report, available on-line at http://reports.weforum.org/global-information-technology-report-2015/network-readiness-index/.

Samoilenko, S. and Osei-Bryson, K.M. (2007). Increasing the Discriminatory Power of DEA in the Presence of the Sample Heterogeneity with Cluster Analysis and Decision Trees, Expert Systems with Applications, 34, 2, 1568-1581.

Samoilenko, S. and Osei-Bryson, K.M. (2014). Formulation of Context-Dependent and Target-Specific Strategies of the Impacts of ICT on Development, in Proceedings of the 7th Annual SIG GlobDev Pre-ICIS Workshop ICT in Global Development; Auckland, New Zealand, December 14, 2014.